Supporting information for: Development of CHARMM-compatible force-field parameters for cobalamin and related cofactors from quantum mechanical calculations

Anna Pavlova,[†] Jerry M. Parks,[*,‡] and James C. Gumbart[*,†,¶]

[†]*School of Physics, Georgia Institute of Technology, Atlanta, GA 30332*

[‡]*Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA*

[¶]*School of Chemistry and Biochemistry, Parker H. Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA 30332*

E-mail: parksjm@ornl.gov; gumbart@physics.gatech.edu

Assignment of LJ parameters for the atoms of the corrin ring by analogy.

See Table S1 for the assignment of the parameters for the heavy atoms. For the hydrogens bound to C15, C10, and C35 we used the atom type HGR61, corresponding to the aromatic hydrogens in the 6-membered benzene ring. For all other hydrogen atoms we used atom types HGA3, HGA2 and HGA1 for the hydrogens bound in CH_3, CH_2 and CH, respectively.

Table S1: Assignment of the LJ parameters for the heavy atoms of the corrin ring of our small B_{12} model by analogy. See Scheme 1B for the PDB atom numbering. For our complete B_{12} model, we used CG3C51 atom types for atoms C2-3, C7-8, C12-13, C17-18 instead of CG3C52. Note, the LJ parameters for C and CH in 5-membered ring are presently same in the CGenFF force field.

Atom Number	Assigned Atom Type	CGenFF atom type
N21-N24	imidazole nitrogen	NG2R50
C4-6, C9-11, C14-16	6-membered ring aromatic carbon	CG2R61
C2-3, C7-8, C12-13, C17-18	CH_2 5-membered ring	CG3C52

Quantum Mechanical Calculations

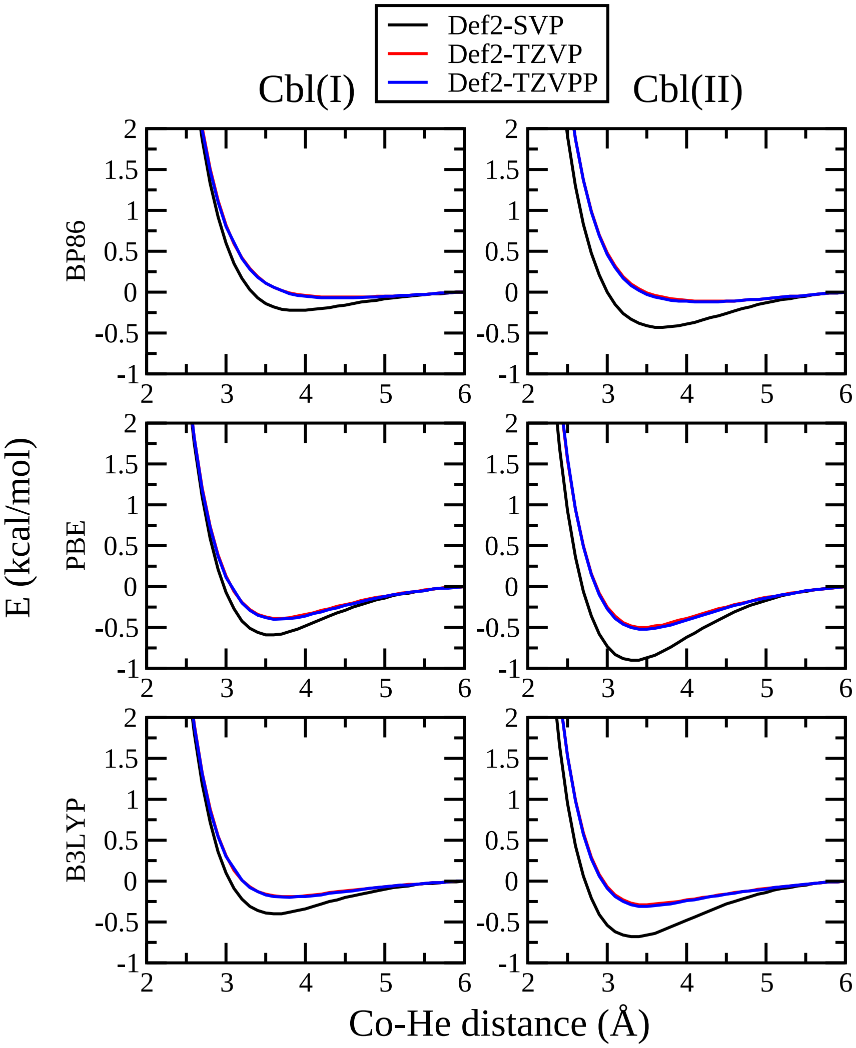

Figure S1: Basis set convergence for Co–He interaction energies for the three considered DFT functionals. The D3 correction was used for all cases.

Fitting of partial atomic charges

Table S2: Partial atomic charges for selected atoms obtained using only RESP for our MeCbl(III) and Cbl(II) models. See Scheme 1B for the atom numbering.

	State	
Atom	MeCbl(III)	Cbl(II)
Co	-0.640	0.044
N21,N24	0.151	-0.076
N23,N22	0.193	-0.155

Table S3: The partial atomic charges for selected atoms when NPA charges were used for cobalt, while RESP was used for all other heavy atoms. See Scheme 1B for the atom numbering.

	State		
Atom	MeCbl(III)	Cbl(II)	Cbl(I)
N21,N24	-0.156	-0.328	-0.42
N23,N22	-0.285	-0.484	-0.592

Table S4: Comparison of NPA-determined charges for selected atoms between MeCbl(III) and AdoCbl(III) models. See Scheme 1B for the atom numbering.

Atom	State	
	MeCbl(III)	AdoCbl(III)
Co	0.382	0.410
N21,N24	-0.403	-0.400
N23,N22	-0.444	-0.455

Fitting of bond and angle parameters

Table S5: Parameters for selected bonds in the MeCbl(III), Cbl(II), and Cbl(I) small models. All K_b values are in kcal mol^{-1} Å^{-2} and all b_0 values are in Å. NP indicates that the bond is not present in the model. The complete set of bond parameters is provided in the Supporting Information.

	B$_{12}$ State					
	MeCbl(III)		Cbl(II)		Cbl(I)	
Bond	K_b	b_0	K_b	b_0	K_b	b_0
Co–N21	61.11	1.84	59.42	1.85	30.732	1.78
Co–N22	56.37	2.00	50.3	1.89	48.68	1.79
Co–N23	55.02	2.00	60.2	1.91	65.22	1.83
Co–N24	55.47	1.86	58.0	1.85	39.26	1.80
Co–N3B	225.84	2.19	122.7	2.20	NP	NP

Table S6: Parameters for selected angles in the MeCbl(III), Cbl(II), and Cbl(I) small models. All K_θ values are in kcal mol^{-1} deg.$^{-2}$ and all θ_0 are in deg. NP indicates that the angle is not present in the model. See the parameter file in Supporting Information for all optimized angle parameters.

	B$_{12}$ State					
	MeCbl(III)		Cbl(II)		Cbl(I)	
Angle	K_θ	θ_0	K_θ	θ_0	K_θ	θ_0
N22–Co–N21	175.01	92.31	27.51	91.52	29.59	90.43
N24–Co–N21	224.96	84.91	53.01	84.51	141.99	84.53
N23–Co–N22	219.19	92.91	78.19	93.60	109.01	92.59
N24–Co–N23	210.24	91.35	56.37	92.46	77.34	92.17
N21–Co–N3B	91.70	88.32	66.67	92.88	NP	NP
N22–Co–N3B	98.85	90.96	53.52	95.92	NP	NP
N23–Co–N3B	222.42	88.62	64.01	95.37	NP	NP
N24–Co–N3B	91.20	91.91	51.19	96.48	NP	NP

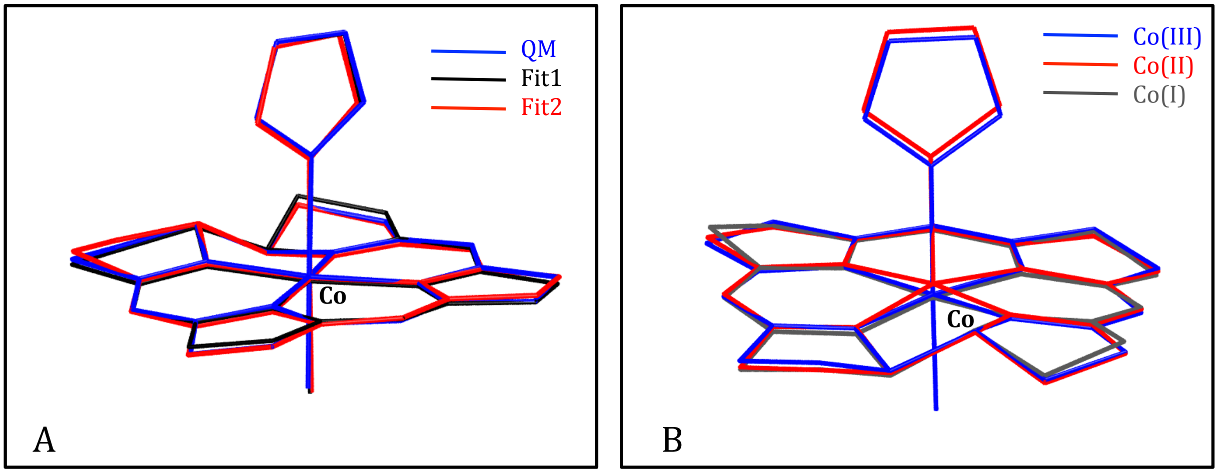

Figure S2: Comparison of geometries. (A) Comparison of aligned minimized geometries of our MeCbl(III) model. Hydrogens are omitted for clarity. QM-minimized geometry is shown in blue, while the geometry obtained with Fit1 parameters (where only two different atom types were used for N21-N24) is shown in black. Finally, the geometry obtained with Fit2 parameters (four different atom types for atoms N21-N24) is shown in red. The Fit1 and Fit2 geometries are shown before the fitting of the dihedral parameters. (B) Comparison of aligned QM-minimized geometries for three of our small B12 models: MeCbl(III) (blue), Cbl(II) (red) and Cbl(I) (black). Hydrogens are omitted for clarity.

Fitting of dihedral parameters

The following dihedrals were refitted from the ones obtained with the MeCbl(III) model or were added. For the Cbl(II) model, we refitted the dihedrals for N21-C1-C19-C18 and N24-C19-C1-C2. For the Cbl(I) model we refitted the dihedrals C6-C7-C8-C9, C11-C12-C13-C14, N23-C11-C12-C13, N23-C14-C13-C12, N22-C9-C8-C7 and N22-C6-C7-C8. The dihedrals for C1-C2-C3-C4 and C26-C17-C18-C19 were refitted using the Con2 model; these new dihedrals were used for all full cobalamin models. The dihedrals C6-C7-C8-C9 and C11-C12-C13-C14 were added to the dihedral fit of the ExtBC model and to all full cobalamin models except Cbl(I).

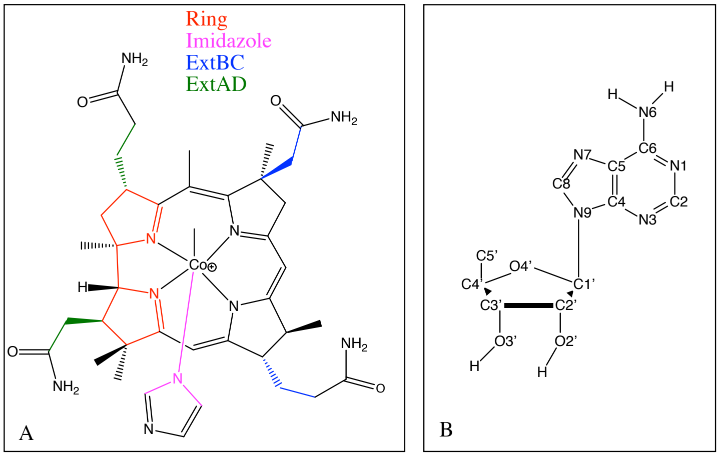

Figure S3: (A) Definition of the dihedral groups - imidazole, ring, ExtBC and ExtAD - that were explicitly optimized here. The different colors show the bonds around which the dihedral scans were performed. (B) PDB atom numbering for the 5-adenosyl ligand.

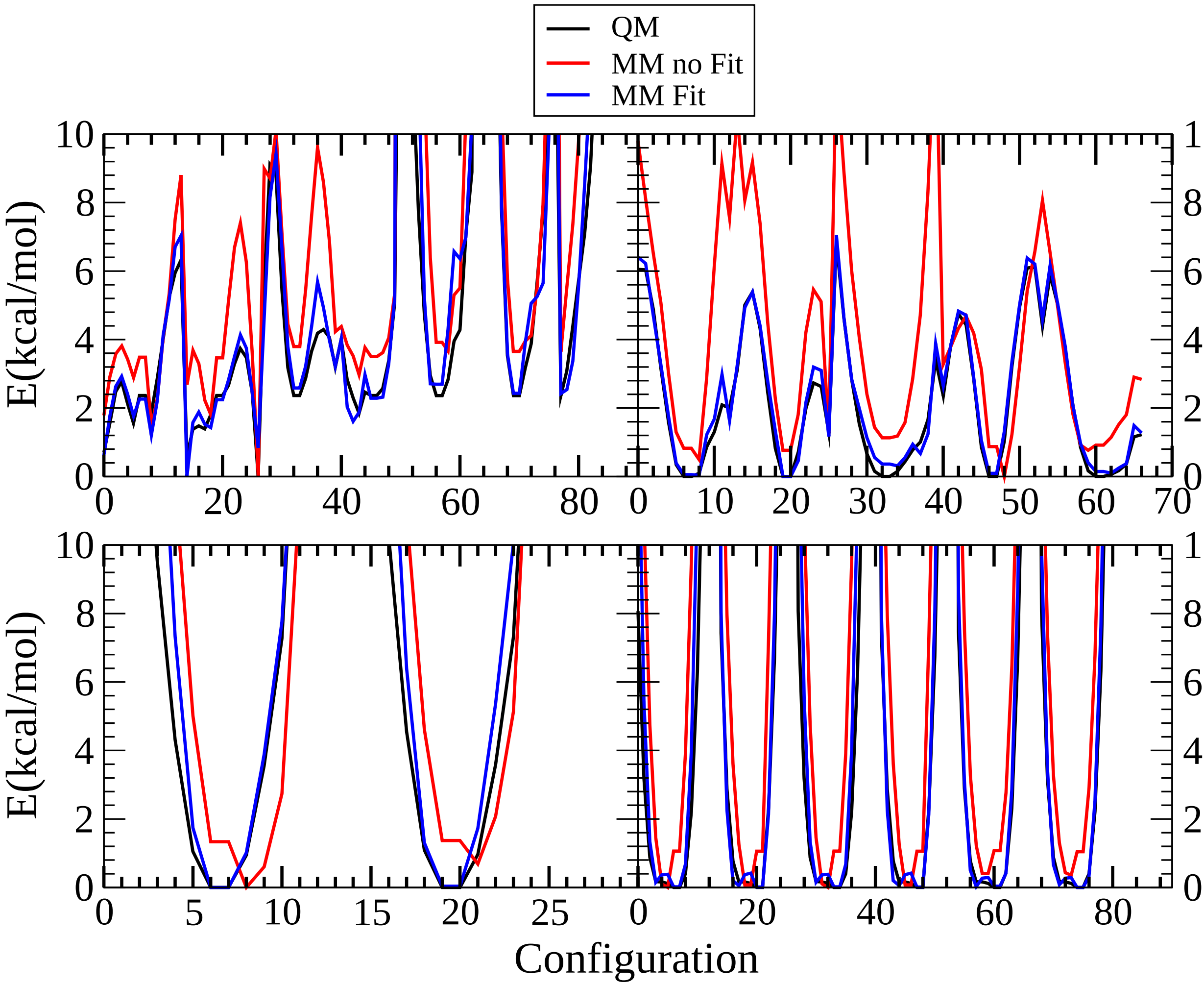

Figure S4: Comparison of the QM-PES to the MM-PES before (no Fit) and after (Fit) fitting of the dihedral constants for several dihedral scans. Top left and top right graphs show the dihedral scans for ExtBC and ExtAD dihedrals, respectively. Bottom left and right graphs shown the dihedral scans for refitted dihedrals for Cbl(II) and Cbl(I) models, respectively.

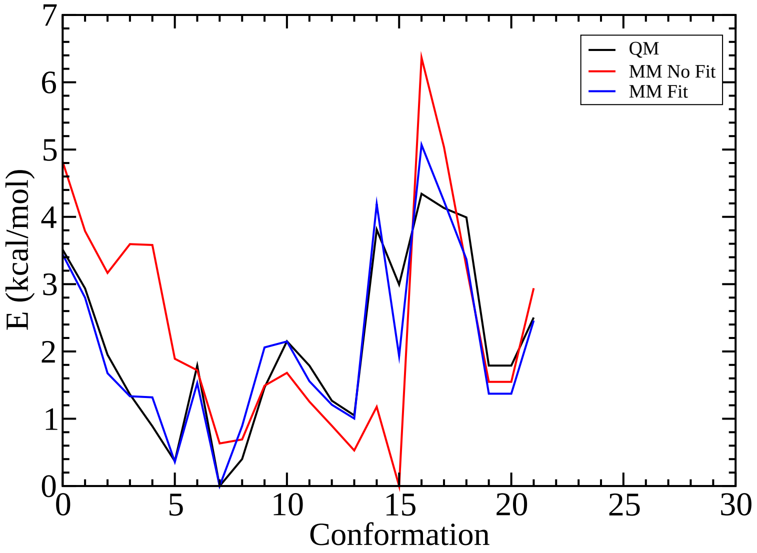

Figure S5: Comparison of the QM-PES to the MM-PES before (no Fit) and after (Fit) fitting of the dihedral parameters for Ado dihedrals.

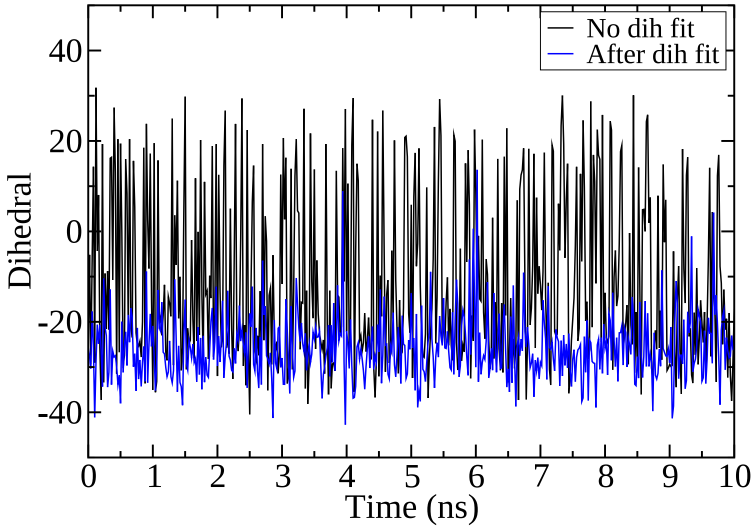

Figure S6: Fluctuations of the C1-C2-C3-C4 dihedral before and after fitting of the ring dihedrals.

Validation

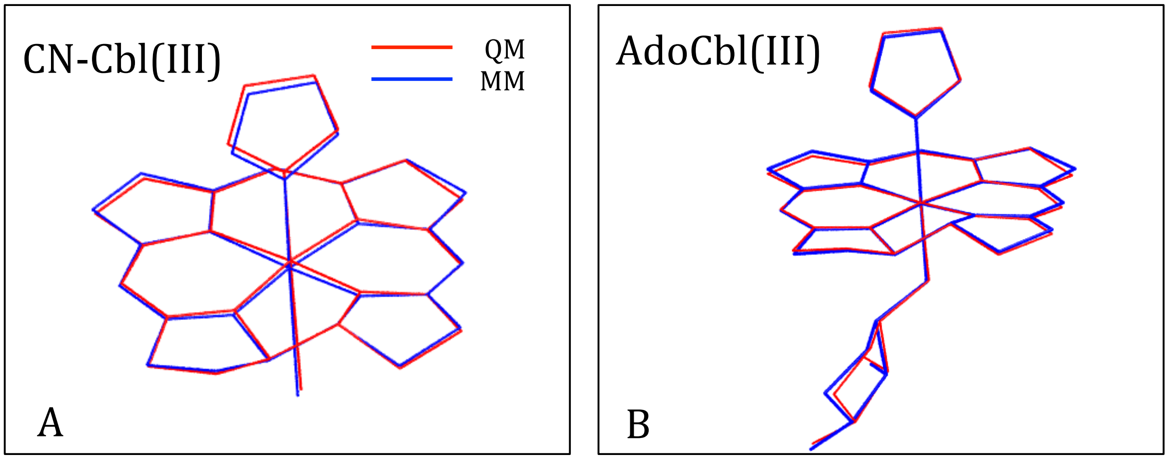

Figure S7: Comparison of QM (red) and MM (blue) geometries for our small (A) CN-Cbl(III) and (B) AdoCbl(III) models. Hydrogens are omitted for clarity.

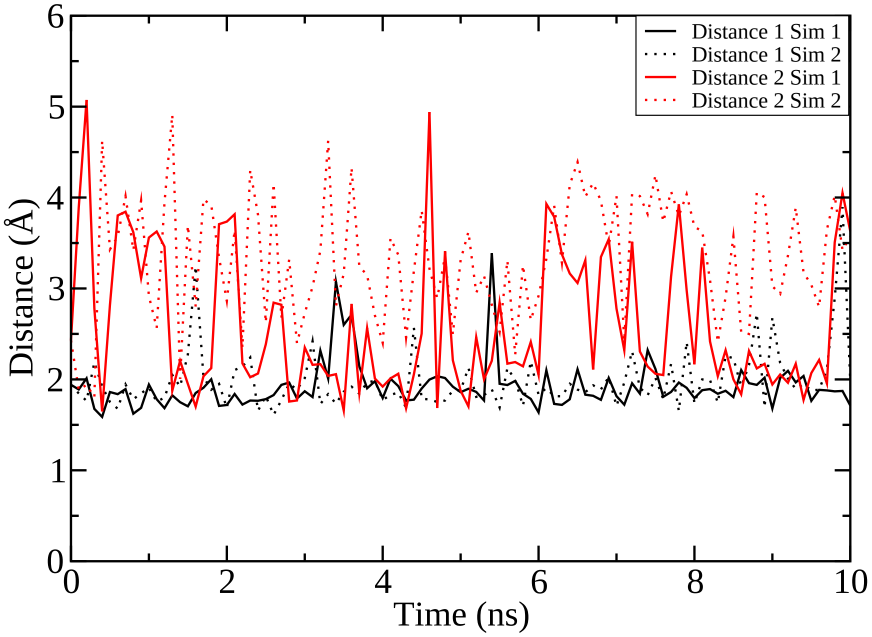

Figure S8: Distances 1 and 2 in Chain B during our MD simulations of CarH (PDB 5C8A). See Figure 4 and Table 8 for the distances in Chain A and distance definitions, respectively.

Fitting of improper force constants

We investigated how increasing the improper force constants changed the resulting fold angles for aqueous MetCbl(III) and Cbl(II) (S7). All improper constants containing nitrogens N21-N24 were increased simultaneously to the stated value in the table. It was found that for values between 40 and 150 kcal mol^{-1} deg^{-2} the fold angle always decreased for MetCbl(III). For Cbl(II) the angle decreased for force constants between 40 and 100 kcal mol^{-1} deg^{-2}; however, almost no difference was observed when the constant was increased from 100 to 150 kcal mol^{-1} deg^{-2}. Therefore, we did not consider values higher than 100 kcal mol^{-1} deg^{-2}. Table S8 shows averaged fold angles during MD simulations for all parametrized forms of B$_{12}$ when the force constant is 80, 90 and 100 kcal mol^{-1} deg^{-2}. The best agreement with the target experimental data was obtained when the the force constant was 100 kcal mol^{-1} deg^{-2}. Hence, this value was used in the final parameter file and MD simulations (MD$_{\mathrm{Fit}}$).

Table S7: Changes in the folding angle for MeCbl(III) and Cbl(II) with increasing improper force constants. The fold angle is given in deg. and improper force constants are given in kcal mol^{-1} deg^{-2}.

System	Force constant				Exp
	40	70	100	150	
MeCbl(III)	17.1	16.0	13.5	11.8	14.8
Cbl(II)	22.0	17.6	14.0	13.6	11.5

Table S8: Fold-angle averages for solvated cobalamins in MD simulations with different improper force constants compared to the experimental angles. Fold angle is given in deg. and improper force constants are given in kcal mol^{-1} deg^{-2}. NA means that experimental data was not available.

System	Force constant				Exp
	Init	80	90	100	
MeCbl(III)	18.4	14.5	13.6	13.5	14.8
CN-Cbl(III)	18.1	12.4.	13.3	12.1	14.1
AdoCbl(III)	15.9	14.5	12.9	10.6	10.0
Cbl(II)	26.5	18.3	16.4	14.0	11.5
Cbl(I)	21.1	11.4	12.3	11.3	NA

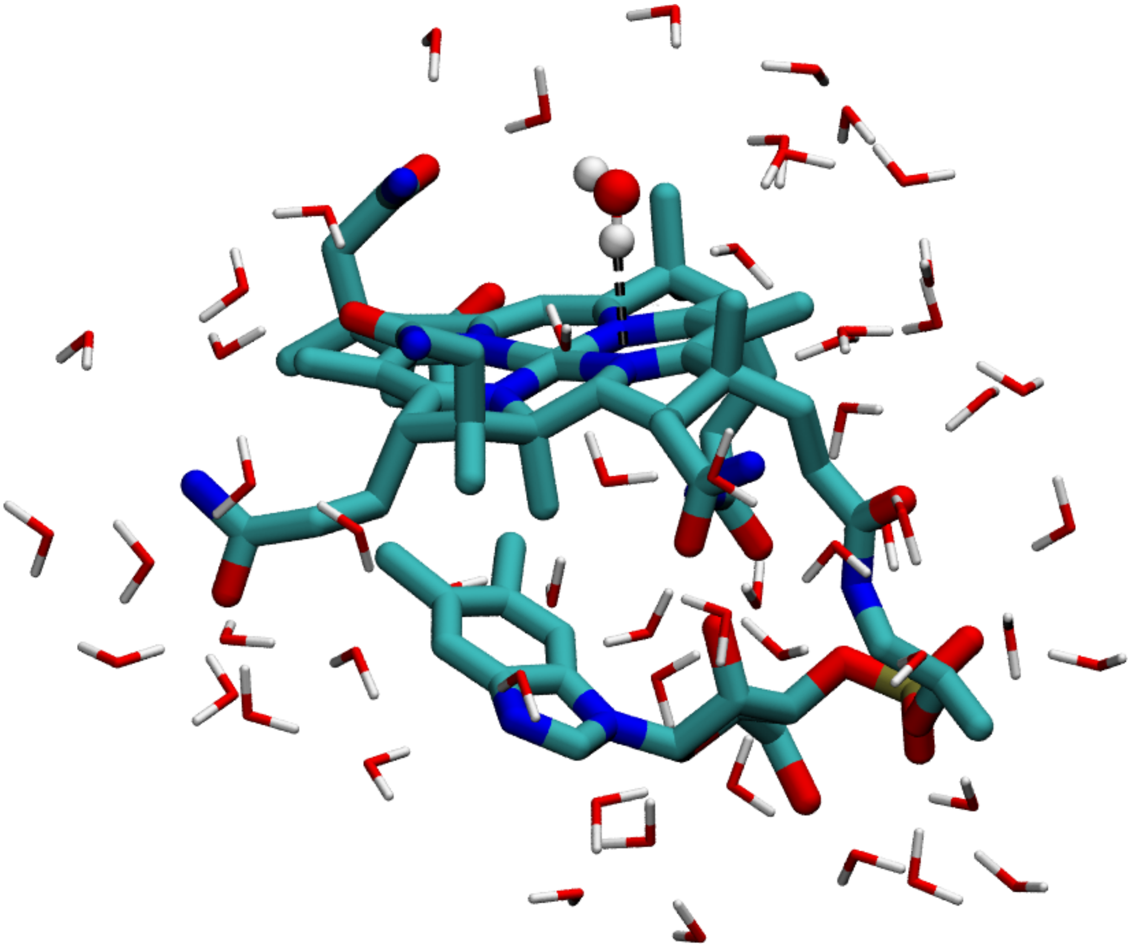

Figure S9: Snapshot of Co(I) cobalamin in water from our simulations, illustrating its interactions with water. Hydrogen atoms of cobalamin are omitted for clarity; water molecules within 3 Å of cobalamin are displayed. The water molecule that forms a hydrogen bond with N24 in the corrin ring is enlarged.

Instructions for using provided B12 topologies with the psfgen plugin in VMD

In addition to the files provided here, the latest CGenFF parameter and topology files are also needed. The provided topology file contains three different residues for cobalamin, corresponding to the three possible oxidation states of the cobalt atom (CO3, CO2 and CO1). Typically, cobalamin has resname B12 in the crystal structure, requiring one to change it in psfgen by using the statement

```
pdbalias residue B12 CO*
```

where * corresponds to the desired oxidation state, 1, 2, or 3. Cyanide is the default β ligand for the Co(III) state. Use patches CH3 and ADE to change this ligand to Met or Ado, respectively. Note, the ADE patch treats the adenosyl ligand as a separate residue with resname 5AD and assumes the standard PDB atom names for the adenine and the deoxyribose sugar. You might need to use the pdbalias command to ensure that the residue and atom names are correct in your segment (see Figure S3B). DMB is the default α ligand; however, for CO(III) and Co(II) states histidine can be bound to the cobalt atom instead by applying a patch. The patch depends on the cobalt oxidation state and the β ligand: HION3 for CN-Cbl(III), HION3A for MeCbl(III) and AdoCbl(III) and HION2 for Cbl(II). When using the patch, specify the resid of cobalamin first, followed by the resid for the histidine.

Note that the ADE or CH3 patches need to be used before the HION patches. For example, in PDB 5C8A, cobalamin is in the AdoCbl(III) state, so both ADE and HION3A patches need to be applied to ensure the correct axial ligands (Ado and His) are used.

In order to simulate norpseudo-B1$_{12}$ and Factor III$_{\mathrm{m}}$, patches PSE and FIII, respectively, should be used.